Adjoint Sensitivity Analysis of Atmospheric Dynamics: Application to the Case of Multiple Observables

EUGENE A. USTINOV

Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California

(Manuscript received 6 March 2000, in final form 25 May 2001)

ABSTRACT

The matrix approach to the adjoint sensitivity analysis of atmospheric models with multiple observables is presented. The approach is developed as a straightforward generalization of the scalar case originally devised by Marchuk in the 1960s for applications in atmospheric remote sensing based on the use of the adjoint equation of radiative transfer. According to the commonly accepted viewpoint, the solution of the adjoint problem corresponding to the forward problem formulated with respect to an n vector of variables, is also an n vector. It is shown that in the general case of m observables this adjoint solution should have the form of an $n \times m$ matrix. Then, the $m \times N$ matrix of sensitivities of m observables to N model parameters can be directly evaluated from the single matrix adjoint solution rather than from multiple vector adjoint solutions computed for each of m observables. Potentially, this can provide appreciable savings of computer time. A general operator–matrix presentation of the approach is given and its application to the sensitivity analysis of a simple zero-dimensional radiative balance model with two field variables and two observables is considered. The results are validated by numerical experiments.

1. Introduction

Currently, there exists a wide variety of models of atmospheric dynamics (see, e.g., Trenberth 1992). The simplest zero-dimensional models describe the spatial averages of the atmospheric field variables. The one-dimensional (1D) models deal with vertical profiles of horizontally averaged variables or with latitudinal cross sections of variables averaged vertically and zonally. The 2D and 3D models add capability to cover the latitudinal and longitudinal variations. There is also a great variety of timescales used, starting from the models of paleoclimate involving the geological timescales and proceeding toward the models of modern climate, and models of interannual, seasonal, synoptic, and diurnal variations, which are used to forecast the long-term and short-term weather phenomena. The most sophisticated models involve integration of the full system of primitive equations of atmospheric motion.

Notwithstanding this diversity of models of atmospheric dynamics, all of them have some common general features. There are a number of *atmospheric variables* that quantify the simulated state of the atmosphere, and there are also a number of *atmospheric parameters* that specify the atmospheric model itself. In general, the models are described by nonlinear differential equations with initial value conditions and/or boundary conditions. As any model has to be eventually tested against observations, an important component of its mathematical description is also a set of *observable results* (*observables*), which are specified by the procedure of their computation from atmospheric variables. These observables can represent a snapshot of atmospheric variables at a given instant, temporal behavior of the observables at given location(s), some space and/or time averages, or some combination of the above.

For a given model, the observables are dependent on the model parameters, and the capability to estimate the sensitivity of observables to these parameters is as important as the capability to compute the observables themselves. Studies of the radiative forcing due to variations of the content, spatial distribution, and temporal behavior of greenhouse atmospheric gases and of aerosol present just one example of where such sensitivity analysis is important. Computation of sensitivities by plain variations of atmospheric parameters may become impractical as the number of parameters increases. If spatial distribution of atmospheric parameters has to be taken into account, then variation of them, layer by layer, grid cell by grid cell, can easily result in insurmountable requirements on the computer resources needed.

An alternative approach to sensitivity analysis in the atmospheric dynamics was suggested in 1970s by Mar-

Corresponding author address: Dr. Eugene A. Ustinov, Jet Propulsion Laboratory, CIT, Mail Stop 169-237, 4800 Oak Grove Drive, Pasadena, CA 91109.
E-mail: Eugene.A.Ustinov@jpl.nasa.gov

chuk (1974, 1975a,b), a decade after he initially introduced this approach in atmospheric remote sensing (Marchuk 1964). Quantitatively, the sensitivities of the observables are derivatives with respect to the atmospheric parameters considered: partial derivatives, as in the zero-dimensional models; and variational derivatives if spatial fields of both observables and atmospheric parameters are involved in models with one or more dimensions. The above plain layer-by-layer, cell-by-cell variation procedure represents essentially a finite-difference approach to evaluation of these derivatives when only numerical solutions of systems of equations of atmospheric dynamics are available. It turns out that the solution of the corresponding system of *adjoint* equations provides an efficient and elegant way to compute sensitivities to the model parameters. A single solution of the adjoint system for a given atmospheric model can be used to compute the sensitivities in a way that is not too dissimilar from that of computing the observables themselves from the solution of the given system of equations of atmospheric dynamics.

The adjoint approach to sensitivity analysis was later developed by Cacuci (1981a,b) in application to the general nonlinear case of a system of equations with initial and boundary conditions. These results were applied to the adjoint sensitivity analysis of the radiative–convective model (Hall et al. 1982; Hall and Cacuci 1983), climate models (Cacuci and Hall 1984), and the general circulation model (Hall 1986). Albeit very general, the mathematical framework developed by Cacuci (1981a,b) was dealing with the case of one observable in the form of a single scalar functional of atmospheric variables and model parameters. Meanwhile, in the case of multiple variables, it is of considerable interest to be able to perform the sensitivity analysis for individual observables without solving the adjoint problem for each of them separately. While the adjoint approach to sensitivity analysis rapidly evolves toward applications to more sophisticated and more realistic models (see, e.g., Kaminski et al. 1999a,b; Vukicevic and Hess 2000; and Li et al. 2000), the ability of direct treatment of multiple observables becomes more and more relevant.

The aim of this paper is to present a straightforward matrix generalization of the formalism of adjoint sensitivity analysis from the scalar case of one variable–one observable devised by Marchuk (1964) for atmospheric remote sensing (see also Ustinov 1991, 1992, 2001), to the case of multiple variables–multiple observables that is applied here to the atmospheric dynamics. In the most general case, the vector of observables can be construed as a nonlinear operator acting on the vectors of atmospheric variables and model parameters. In many practical applications of atmospheric modeling, the vector of observables can be construed as a linear operator representing some weighted averages over space and/or time of separate atmospheric variables or of any linear combination of them. Thus, the procedure of computing the observables involves, in general, a *matrix* transformation; in the case of observables computed from separate atmospheric variables the transformation matrix is diagonal. As shown in section 3, in general case, this matrix is constructed from the right-hand terms of the matrix differential equation and initial condition of the adjoint problem of atmospheric dynamics. This dictates the matrix nature of the adjoint solution in the general case.

2. Nonlinear and linearized forward problems

In the following, we will use different superscripts to distinguish between variables and parameters used in the formulation of three interrelated problems. The variables and parameters without superscripts refer to the basic nonlinear forward problem represented by the system of equations of atmospheric dynamics. The variables and parameters with prime (′) superscript refer to the linearized forward problem based on the equations of atmospheric dynamics, which are linearized in the vicinity of the nonlinear solution and which describe the perturbation of this solution due to perturbations of parameters of the system. Finally, the star (*) superscript denotes the adjoint variables that represent the solution of the linear problem that is adjoint to the linearized forward problem.

Let $\mathbf{X}(t)$ be an n vector of variables $X_j(t)$, ($j = 1, \ldots n$) describing the state of the atmosphere. Let $\mathbf{a}(t)$ be an N vector of model atmospheric parameters, $a_i(t)$ ($i = 1, \ldots N$). The system of nonlinear equations describing evolution of the atmospheric system together with the initial conditions for variables $X_j(t)$ can be written in the form

$$\frac{d\mathbf{X}}{dt} + \mathcal{N}(\mathbf{X}, \mathbf{a}) = 0 \tag{1}$$

$$\mathbf{X}(t_0) = \mathbf{X}_0. \tag{2}$$

Here $\mathcal{N}$ is a nonlinear operator that acts on the state vector $\mathbf{X}$ in a way that depends on the vector of atmospheric parameters, $\mathbf{a}$. Let $\mathbf{X}(t)$ be a solution of this nonlinear model computed for the period of evolution of the atmospheric system for the interval of time from $t = t_0$ to $t = t_1$. Let $\mathbf{R}$ be an m vector of observables R_k, ($k = 1, \ldots m$) obtained using the procedure of observations that is specified by the $n \times m$ observation matrix $\mathsf{W}(t)$ as convolved with the vector of variables $\mathbf{X}(t)$ over the interval of integration $[t_0, t_1]$:

$$\mathbf{R} = \int_{t_0}^{t_1} \mathsf{W}^{\mathrm{T}}(t)\mathbf{X}(t)\, dt. \tag{3}$$

For example, if the vector of observables $\mathbf{R}$ represents the state vector $\mathbf{X}$ at a given instant, t_s, then $m = n$ and

$$\mathsf{W}(t) = \delta(t - t_s)\mathsf{I}, \tag{4}$$

where I is an $n \times n$ identity matrix and $\delta(t)$ is the Dirac

δ function. If the observables represent the state vector averaged over the period from t_{s1} to t_{s2}, then

$$\mathbf{W}(t) = [\theta(t - t_{s1}) - \theta(t - t_{s2})]\mathbf{I}, \tag{5}$$

where $\theta(t)$ is the Heaviside θ function.

If the number of observables m is less than the number of model variables n, then the observation matrix $\mathbf{W}(t)$ has correspondingly a lower number of columns. And if only one observable result is considered, the matrix $\mathbf{W}(t)$ degenerates into the n vector. If, in addition, the time dependence of $\mathbf{W}(t)$ is the same for all its elements, then it can be rewritten in the form of the n vector $\mathbf{d}$ multiplied by a scalar function:

$$\mathbf{W}(t) = W(t)\mathbf{d}. \tag{6}$$

In this particular case, Eq. (3) can be rewritten as [cf. Eq. (2) of Hall and Cacuci (1983)]

$$R = \int_{t_0}^{t_1} W(t)\mathbf{d}^{\mathrm{T}}\mathbf{X}(t)\, dt. \tag{7}$$

The general aim of sensitivity analysis is to evaluate the responses of the vector of observables $\mathbf{R}$ to variations of the vector of parameters $\mathbf{a}(t)$. Let these parameters experience some variations $\delta\mathbf{a}(t)$ in the vicinity of values $\mathbf{a}(t)$ for which the solution of the nonlinear system $\mathbf{X}(t)$ was found. We will refer to $\mathbf{X}(t)$ as to a basic nonlinear solution. Let the nonlinear problem, Eqs. (1) and (2), be linearized in vicinity of $\mathbf{X}(t)$ in the form

$$\frac{d\mathbf{X}'}{dt} + \mathbf{C}(t)\mathbf{X}'(t) = \mathbf{S}_e(t) \tag{8}$$

$$\mathbf{X}'(t_0) = \mathbf{S}_c. \tag{9}$$

Here $\mathbf{C}(t)$ is the $n \times n$ matrix and $\mathbf{S}_e(t)$ is an n vector. Both of them are dependent on the given basic solution $\mathbf{X}(t)$. If, in addition to the atmospheric parameters, initial conditions are also varied, then $\mathbf{S}_c \neq \mathbf{0}$. Subscripts e and c stand for "equation" and "(initial) condition," respectively.

In the next section, we will need to combine the matrix differential equation and initial condition of the linearized forward problem, Eqs. (8) and (9) into a single linear operator equation using the scheme developed in Ustinov (2001) for the scalar case of atmospheric radiative transfer. For this purpose, we represent Eqs. (8) and (9) in a general form of two linear operator equations:

$$\mathcal{L}_e\mathbf{X}' = \mathbf{S}_e \tag{10}$$

$$\mathcal{L}_c\mathbf{X}' = \mathbf{S}_c \quad \text{at } t = t_0, \tag{11}$$

where linear operators $\mathcal{L}_e$ and $\mathcal{L}_c$ represent linear operations on $\mathbf{X}'$ in left sides of the equation and initial condition [Eqs. (8), (9)].

In order to formulate the matrix adjoint problem in the next section we also need a linearized version of Eq. (3) to obtain the vector of corresponding perturbations of observables $\mathbf{R}'$ as expressed through $\mathbf{X}'$:

$$\mathbf{R}' = \int_{t_0}^{t_1} \mathbf{W}^{\mathrm{T}}(t)\mathbf{X}'(t)\, dt. \tag{12}$$

In the development of the adjoint matrix formalism in the next section we will use the following definition. Let $\mathbf{A}(t)$ be an $n \times m$ matrix function and $\mathbf{B}(t)$ be an n-vector function. Their inner product $(\mathbf{A}, \mathbf{B})$ is defined as an m vector in the form

$$(\mathbf{A}, \mathbf{B}) = \int_{t_0}^{t_1} \mathbf{A}^{\mathrm{T}}(t)\mathbf{B}(t)\, dt. \tag{13}$$

For the elements of the vector $(\mathbf{A}, \mathbf{B})$ we have

$$(\mathbf{A}, \mathbf{B})_k = \sum_{j=1}^{n} (A_{jk}, B_j) \qquad (k = 1, \ldots m), \tag{14}$$

where (A_{jk}, B_j) $(j = 1, \ldots n;\ k = 1, \ldots m)$ are the inner products of corresponding scalar functions $A_{jk}(t)$ and $B_j(t)$. If $m = n = 1$, then Eq. (13) becomes a definition of an inner product of two scalar functions $A(t)$ and $B(t)$. Using the definition Eq. (13), we can rewrite Eq. (12) in the form

$$\mathbf{R}' = (\mathbf{W}, \mathbf{X}'). \tag{15}$$

In the particular case of one observable, Eq. (7) we have

$$\mathbf{R}' = (W\mathbf{d}, \mathbf{X}'). \tag{16}$$

3. Adjoint operator and adjoint problem

In this section we will construct the adjoint operator and adjoint problem corresponding to the linearized forward problem [Eqs. (10), (11)] coupled with the procedure [Eq. (15)] of computation of the (linearized) observables from the linearized solution. These equations are used to formulate the corresponding adjoint problem in the form

$$\mathcal{L}_e^*\mathbf{X}^* = \mathbf{W}_e \tag{17}$$

$$\mathcal{L}_c^*\mathbf{X}^* = \mathbf{W}_c \quad \text{at } t = t_1. \tag{18}$$

Following the scheme developed in Ustinov (2001), first we will combine the differential equation and boundary condition of the forward problem, Eqs. (10) and (11) into a single operator equation,

$$\mathcal{L}\mathbf{X}' = \mathbf{S}, \tag{19}$$

and then we will find an operator $\mathcal{L}^*$, adjoint to $\mathcal{L}$. As was demonstrated in Ustinov (2001) for the scalar case, this operator will naturally split into the operators $\mathcal{L}_e^*$ and $\mathcal{L}_c^*$, corresponding to the differential equation of the adjoint problem and to a condition that is imposed here at the end of the time interval t_1. Then, based on the derived form of the operator $\mathcal{L}^*$ and on the given procedure [Eq. (15)] of computation of observables $\mathbf{R}'$, we derive the matrices $\mathbf{W}_e$ and $\mathbf{W}_c$ in the right sides of Eqs. (17) and (18).

To combine Eqs. (10) and (11) into a single linear operator equation in the form of Eq. (19), we observe that Eq. (10) holds for the whole time interval $[t_0, t_1]$ while Eq. (11) holds for the instant $t = t_0$ only. After multiplying Eq. (11) by the time-weighting factor $\delta(t - t_0)$ we can add it to Eq. (11) to obtain a single operator equation:

$$[\mathcal{L}_e + \delta(t - t_0)\mathcal{L}_c]\mathbf{X}' = \mathbf{S}_e + \delta(t - t_0)\mathbf{S}_c. \quad (20)$$

Comparing Eqs. (20) and (19) we obtain the operator $\mathcal{L}$ and the right-hand term $\mathbf{S}$ of the forward problem, [Eq. (19)] in the form

$$\mathcal{L} = \mathcal{L}_e + \delta(t - t_0)\mathcal{L}_e \quad (21)$$

$$\mathbf{S} = \mathbf{S}_e + \delta(t - t_0)\mathbf{S}_c. \quad (22)$$

From Eqs. (8) and (9) we have

$$\mathcal{L}_e = \frac{d}{dt} + \mathbf{C}(t) \quad (23)$$

$$\mathcal{L}_c = \mathbf{I}, \quad (24)$$

where $\mathbf{I}$ is an identity matrix. After substitution into Eq. (21) we obtain

$$\mathcal{L} = \frac{d}{dt} + \mathbf{C}(t) + \delta(t - t_0)\mathbf{I}. \quad (25)$$

As in the scalar case, to obtain the operator $\mathcal{L}^*$, which is adjoint to the operator $\mathcal{L}$ [Eq. (25)], we use the formal definition of $\mathcal{L}^*$ as an operator satisfying the identity

$$(\mathbf{X}^*, \mathcal{L}\mathbf{X}') = (\mathcal{L}^*\mathbf{X}^*, \mathbf{X}') \quad (26)$$

for arbitrary functions $\mathbf{X}'(t)$ and $\mathbf{X}^*(t)$, which can form the inner product $(\mathbf{X}^*, \mathbf{X}')$ as defined by Eq. (13). In other words, we demand that the adjoint operator $\mathcal{L}^*$ satisfies the equality

$$\int_{t_0}^{t_1} \mathbf{X}^{*\mathrm{T}}\mathcal{L}\mathbf{X}'\, dt = \int_{t_0}^{t_1} (\mathcal{L}^*\mathbf{X}^*)^{\mathrm{T}}\, \mathbf{X}'\, dt. \quad (27)$$

Substituting operator $\mathcal{L}$ as defined by Eq. (25) into the left side of Eq. (27) we have

$$(\mathbf{X}^*, \mathcal{L}\mathbf{X}') = \int_{t_0}^{t_1} \left[\mathbf{X}^{*\mathrm{T}}(t)\frac{d\mathbf{X}'}{dt} + \mathbf{X}^{*\mathrm{T}}(t)\mathbf{C}(t)\mathbf{X}'(t)\right] dt$$
$$+ \mathbf{X}^{*\mathrm{T}}(t_0)\mathbf{X}'(t_0). \quad (28)$$

Performing the integration by parts in Eq. (28) we obtain

$$(\mathbf{X}^*, \mathcal{L}\mathbf{X}') = \int_{t_0}^{t_1} \left[-\frac{d\mathbf{X}^*}{dt} + \mathbf{C}^{\mathrm{T}}(t)\mathbf{X}^*(t)\right]^{\mathrm{T}} \mathbf{X}'(t)\, dt$$
$$+ \mathbf{X}^{*\mathrm{T}}(t_1)\mathbf{X}'(t_1). \quad (29)$$

The right-hand side of Eq. (29) can be represented in the form of the right-hand side of Eq. (27) if we let

$$\mathcal{L}^* = -\frac{d}{dt} + \mathbf{C}^{\mathrm{T}}(t) + \delta(t_1 - t)\mathbf{I}. \quad (30)$$

This means that the operator $\mathcal{L}^*$ of the adjoint problem corresponding to the forward problem [Eqs. (10), (11)] can be represented in the form of a sum of two components:

$$\mathcal{L}^* = \mathcal{L}_e^* + \delta(t_1 - t)\mathcal{L}_c^*, \quad (31)$$

where

$$\mathcal{L}_e^* = -\frac{d}{dt} + \mathbf{C}^{\mathrm{T}}(t); \quad \text{and} \quad (32)$$

$$\mathcal{L}_c^* = \mathbf{I}. \quad (33)$$

From Eq. (31) we see that the operator $\mathcal{L}_e^*$ is acting on the whole time interval $[t_0, t_1]$ while operator $\mathcal{L}_c^*$ is acting only at the final instant t_1 of this interval. Thus, the adjoint problem corresponding to Eqs. (10) and (11) can be represented in the form of Eqs. (17) and (18). Substituting the obtained expressions for $\mathcal{L}_e^*$ and $\mathcal{L}_c^*$ into Eqs. (17) and (18) we have

$$-\frac{d\mathbf{X}^*}{dt} + \mathbf{C}^{\mathrm{T}}(t)\mathbf{X}^*(t) = \mathbf{W}_e(t) \quad (34)$$

$$\mathbf{X}^*(t_1) = \mathbf{W}_c. \quad (35)$$

The matrix right-hand terms, $\mathbf{W}_e$ and $\mathbf{W}_c$, are defined using the relation

$$\mathbf{W} = \mathbf{W}_e + \delta(t_1 - t)\mathbf{W}_c, \quad (36)$$

which is obtained by multiplying Eq. (33) by $\delta(t_1 - t)$, adding it to Eq. (32) and comparing with the general form of the adjoint problem

$$\mathcal{L}^*\mathbf{X}^* = \mathbf{W}, \quad (37)$$

with the matrix right-hand term $\mathbf{W}$. If, for example, the vector of observables $\mathbf{R}'$ [Eq. (12) is defined by some weighted averages over the interval $[t_0, t_1]$, then $\mathbf{W}_c \equiv 0$ and $\mathbf{W}(t) = \mathbf{W}_e(t)$. If, on the other hand, $\mathbf{R}'$ is defined by the state vector $\mathbf{X}'$ at the final instant t_1, then $\mathbf{W}_e(t) \equiv 0$ and $\mathbf{W}(t) = \delta(t_1 - t)\mathbf{W}_c$.

It should be emphasized that the adjoint problem [Eqs. (34), (35)] is formulated with the *matrix* right-hand terms, $\mathbf{W}_e(t)$ and $\mathbf{W}_c$, which have the number of columns corresponding to the number of observables. The adjoint solution is also a matrix with the same dimensions. The matrix $\mathbf{C}(t)$ used in the differential equation, Eq. (34) has to be computed only once at each time step of integration, independent of the number of observables. In other words, the matrix differential equation Eq. (34) is integrated for all columns of the adjoint solution simultaneously. This is favorable as compared to separate integration of Eq. (34) for each observable because the matrix $\mathbf{C}^{\mathrm{T}}(t)$ is not dependent on individual observables. Therefore, in the computer program, it can be evaluated outside of the loop over individual observables, resulting in corresponding savings of computing time.

4. Applications to the sensitivity analysis

The solution $\mathbf{X}^*$ of the adjoint problem [Eq. (37)] is instrumental in two respects. First, it provides an alternative way to compute the vector of observables, in addition to that provided by Eq. (15):

$$\mathbf{R}' = (\mathbf{X}^*, \mathbf{S}). \tag{38}$$

Second, and more importantly, the adjoint solution $\mathbf{X}^*$ makes it possible to directly express the variation of the vector of observables $\mathbf{R}$ through the variation of the operator $\mathcal{L}$ and the right-hand term $\mathbf{S}$ of the forward problem, Eq. (19):

$$\delta\mathbf{R} = (\mathbf{X}^*, \delta\mathbf{S} - \delta\mathcal{L}\mathbf{X}'). \tag{39}$$

Since both operator $\mathcal{L}$ and right-hand term $\mathbf{S}$ are expressed through the parameters $\mathbf{a}_i(t)$ $(i = 1, \ldots N)$, we can obtain the explicit expression for the $m \times n$ sensitivity matrix of the vector of observables $\mathbf{R}$ to the vector of model parameters $\mathbf{a}(t)$. It has the form of a variational derivative $\delta\mathbf{R}/\delta\mathbf{a}(t)$ and enters the relation between variations $\delta\mathbf{R}$ and $\delta\mathbf{a}(t)$:

$$\delta\mathbf{R} = \int_{t_0}^{t_1} \frac{\delta\mathbf{R}}{\delta\mathbf{a}(t)}\delta\mathbf{a}(t)\,dt. \tag{40}$$

[A brief summary of necessary information on variational derivatives can be found, e.g., in the appendix to Ustinov (2000).] From Eq. (39) we have

$$\frac{\delta\mathbf{R}}{\delta\mathbf{a}(t)} = \left(\mathbf{X}^*, \frac{\delta\mathbf{S}}{\delta\mathbf{a}(t)} - \frac{\delta\mathcal{L}}{\delta\mathbf{a}(t)}\mathbf{X}'\right). \tag{41}$$

Equation (38) can be derived by multiplying the forward problem, Eq. (19), by $\mathbf{X}^*$ and multiplying the adjoint problem, Eq. (37), by $\mathbf{X}'$ to obtain

$$(\mathbf{X}^*, \mathcal{L}\mathbf{X}') = (\mathbf{X}^*, \mathbf{S}) \tag{42}$$

$$(\mathcal{L}^*\mathbf{X}^*, \mathbf{X}') = (\mathbf{W}, \mathbf{X}'). \tag{43}$$

As the left-hand terms of Eqs. (42) and (43) are equal by definition of the adjoint operator $\mathcal{L}^*$, and the right-hand term of Eq. (43) equals $\mathbf{R}'$ [cf. Eq. (15)], we immediately obtain Eq. (38).

To obtain Eq. (39) we consider the perturbed linearized forward problem:

$$(\mathcal{L} + \delta\mathcal{L})(\mathbf{X}' + \delta\mathbf{X}) = (\mathbf{S} + \delta\mathbf{S}). \tag{44}$$

Subtracting from Eq. (44) the linearized forward problem, Eq. (19), and neglecting the second-order term containing $\delta\mathcal{L}\delta\mathbf{X}$ we obtain

$$\mathcal{L}\delta\mathbf{X} + \delta\mathcal{L}\mathbf{X}' = \delta\mathbf{S}. \tag{45}$$

Multiplying Eq. (37) by $\delta\mathbf{X}$ and Eq. (45) by $\mathbf{X}^*$ we have

$$(\mathcal{L}^*\mathbf{X}^*, \delta\mathbf{X}) = (\mathbf{W}, \delta\mathbf{X}) \tag{46}$$

$$(\mathbf{X}^*, \mathcal{L}\delta\mathbf{X}) + (\mathbf{X}^*, \delta\mathcal{L}\mathbf{X}') = (\mathbf{X}^*, \delta\mathbf{S}). \tag{47}$$

Taking variation of Eq. (15),

$$\delta\mathbf{R} = (\mathbf{W}, \delta\mathbf{X}), \tag{48}$$

and applying the definition of the adjoint operator $\mathcal{L}^*$, Eq. (26), to the adjoint solution $\mathbf{X}^*$ and to the variation of forward solution $\delta\mathbf{X}$ we have

$$(\mathbf{X}^*, \mathcal{L}\delta\mathbf{X}) = (\mathcal{L}^*\mathbf{X}^*, \delta\mathbf{X}). \tag{49}$$

Finally, subtracting Eq. (47) from Eq. (46) and using Eqs. (48) and (49) we obtain Eq. (39).

In the particular case when $\mathbf{X}' = \mathbf{0}$, Eq. (39) yields the variation $\mathbf{R}$ in the vicinity of the basic nonlinear solution $\mathbf{X}$ of the forward problem in its initial form, Eqs. (1) and (2). The equation

$$\delta\mathbf{R} = (\mathbf{X}^*, \delta\mathbf{S}) \tag{50}$$

and the expression for the sensitivity matrix, Eq. (41), is reduced to the form

$$\frac{\delta\mathbf{R}}{\delta\mathbf{a}(t)} = \left(\mathbf{X}^*, \frac{\delta\mathbf{S}}{\delta\mathbf{a}(t)}\right). \tag{51}$$

In the next section, the matrix approach to sensitivity analysis developed above is applied to a simple radiative balance model of atmospheric dynamics with two variables and two observables.

5. Sensitivity analysis of a simple radiative balance model

Two components of the state vector in this model are temperature T and cloudiness n:

$$\mathbf{X}(t) = \begin{pmatrix} T(t) \\ n(t) \end{pmatrix}. \tag{52}$$

The model is described by two equations with corresponding initial value conditions for T and n:

$$c_p\frac{dT}{dt} + \left(1 - \frac{n}{2}\right)\sigma T^4 = (1 - nA)E_\odot$$
$$\tau\frac{dn}{dt} + n = \frac{T - T^{(0)}}{\Delta t} \tag{53}$$

$$T(t_0) = T_0, \qquad n(t_0) = n_0. \tag{54}$$

The first equation of the system, Eq. (53), is the radiative balance equation with solar heating dependent on the cloudiness. If cloudiness $n = 0$, then the radiative balance is determined by solar heating $E_\odot$ and thermal cooling σT^4 of the surface. If cloudiness $n = 1$, then the radiative balance is determined by solar heating $(1 - A)E_\odot$ and thermal cooling $\sigma T^4/2$ of opaque clouds with cloud tops at the tropopause. The equation for cloudiness includes a characteristic time constant, τ, and results in a linear dependence of n on T in the equilibrium state.

The 2 vector of observables $\mathbf{R}$ is defined by Eq. (3) with the 2×2 matrix $\mathbf{W}(t)$ specified under assumption that the observables R_1 and R_2 are obtained as some weighted averages of two linear combinations of components X_1 and X_2 of the state vector $\mathbf{X}$ within the integration period $[t_0, t_1]$. Then,

$$\mathbf{W}_e(t) = \mathbf{W}(t) \tag{55}$$

$$\mathbf{W}_c = 0. \tag{56}$$

For the sake of simplicity we assume that the only model parameter to vary here is the albedo of clouds A.

Linearization of the nonlinear forward problem, Eqs. (53) and (54), in the vicinity of some basic nonlinear solution, $\mathbf{X}(t)$, yields

$$c_p\frac{dT'}{dt} + 4\sigma T^3(t)T' + \left(-\frac{\sigma}{2}T^4 + AE_\odot\right)n' = -A'E_\odot(t)$$

$$\tau\frac{dn'}{dt} - \frac{1}{\Delta T}T' + n' = 0 \tag{57}$$

$$T'(t_0) = 0, \qquad n'(t_0) = 0. \tag{58}$$

The system [Eqs. (57), (58)] can be presented in the form of the linearized forward model [Eqs. (8), (9)], where 2×2 matrix $\mathbf{C}(t)$ and 2 vectors $\mathbf{S}_e(t)$ and $\mathbf{S}_c$ have the form

$$\mathbf{C}(t) = \begin{pmatrix} \frac{4}{c_p}\sigma T^3 & \frac{1}{c_p}\left(-\frac{\sigma}{2}T^4 + AE_\odot\right) \\ -\frac{1}{\tau\Delta T} & \frac{1}{\tau} \end{pmatrix} \tag{59}$$

$$\mathbf{S}_e(t) = \begin{pmatrix} -\frac{1}{c_p}A'E_\odot(t) \\ 0 \end{pmatrix} \tag{60}$$

$$\mathbf{S}_c = 0. \tag{61}$$

Corresponding adjoint model can be obtained directly from the general matrix form [Eqs. (34), (35)] by substitution of the matrix $\mathbf{C}(t)$ [Eq. (59)]. Both the solution $\mathbf{X}^*(t)$ and the right-hand terms $\mathbf{W}_e(t)$ and $\mathbf{W}_c$ are 2×2 matrices. For each kth column of these matrices we have ($k = 1, 2$)

$$-\frac{dX^*_{1k}}{dt} + \frac{4}{c_p}\sigma T^3 X^*_{1k} - \frac{1}{\tau\Delta T}X^*_{2k} = (W_e)_{1k}$$

$$-\frac{dX^*_{2k}}{dt} + \frac{1}{c_p}\left(-\frac{\sigma}{2}T^4 + AE_\odot\right)X^*_{1k} + \frac{1}{\tau}X^*_{2k} = (W_e)_{2k} \tag{62}$$

$$X^*_{1k}(t_1) = (W_c)_{1k}, \qquad X^*_{2k}(t_1) = (W_c)_{2k}. \tag{63}$$

To use the variational relation for the vector of observables, $\delta\mathbf{R}$ [Eq. (39)], we have to evaluate the variations of the operator $\mathcal{L}$ [Eq. (25)] and the right-hand term $\mathbf{S}$ [Eq. (22)]. Substituting the matrix $\mathbf{C}(t)$ [Eq. (59)] into Eq. (25) and taking the variation with respect to albedo A, we have

$$\delta\mathcal{L} = \begin{pmatrix} 0 & \frac{E_\odot}{c_p}\delta A \\ 0 & 0 \end{pmatrix}. \tag{64}$$

Substituting vectors $\mathbf{S_e}(t)$ and $\mathbf{S_c}$ [Eqs. (60), (61)] into Eq. (22) and taking the variation with respect to A we have

$$\delta\mathbf{S} = \begin{pmatrix} -\frac{E_\odot}{c_p}\delta A \\ 0 \end{pmatrix}. \tag{65}$$

Evaluating

$$\delta\mathcal{L}\mathbf{X}' = \begin{pmatrix} -\frac{E_\odot}{c_p}n'\delta A \\ 0 \end{pmatrix}, \tag{66}$$

substituting Eqs. (65) and (66) into the expression for $\delta\mathbf{R}$ [Eq. (39)], and performing a matrix multiplication, we have

$$\delta R_k = \int_{t_0}^{t_1} X^*_{k1}(t)\left[-\frac{1}{c_p}E_\odot(t)(1 + n'(t))\right]\delta A(t)\, dt \qquad (k = 1, 2), \tag{67}$$

whence we immediately obtain

$$\frac{\delta R_k}{\delta A(t)} = -\frac{1}{c_p}E_\odot(t)(1 + n'(t))X^*_{k1}(t) \qquad (k = 1, 2). \tag{68}$$

TABLE 1. Input data used in the numerical experiments.

Parameter	Input values
Thermal capacity c_p	10^6 J K^{-1} m^{-2}
Cloudiness time constant τ	10^4 s
Cloudiness parameter $T^{(0)}$	273.16 K
Cloudiness parameter ΔT	100 K
Initial temperature T_0	273.16 K
Initial cloudiness n_0	0
Day duration D	8.64×10^4 s
Initial moment of integration t_0	0 s
Final moment of integration t_1	$t_0 + 7D$
Albedo A	0.8
Insolation:	
Variable	max $[E_{\odot 0}\cos 2\pi(t - t_0)/D, 0]$
Constant	$E_{\odot 0}/\pi$
Solar constant $E_{\odot 0}$	1400 W m^{-2}
Observation matrix $\mathbf{W}(t)$	$[\theta(t - t_2) - \theta(t - t_1)]\mathbf{I}$
Initial moment of observations t_2	$t_1 - D$

6. Numerical experiments

The summary of the input data used in the numerical experiments is presented in Table 1. The value of c_p was intentionally taken much less than the typical value for

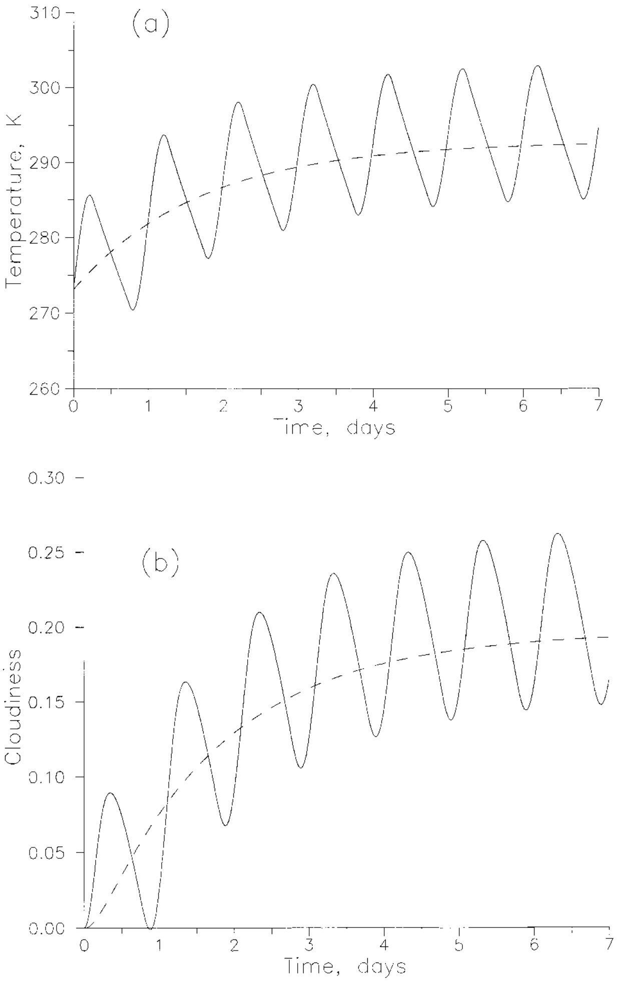

FIG. 1. Basic solution of the nonlinear forward model for the variable (solid line) and constant (dashed line) insolation: (a) temperature; (b) cloudiness. The constant insolation corresponds to the diurnal average of the variable insolation.

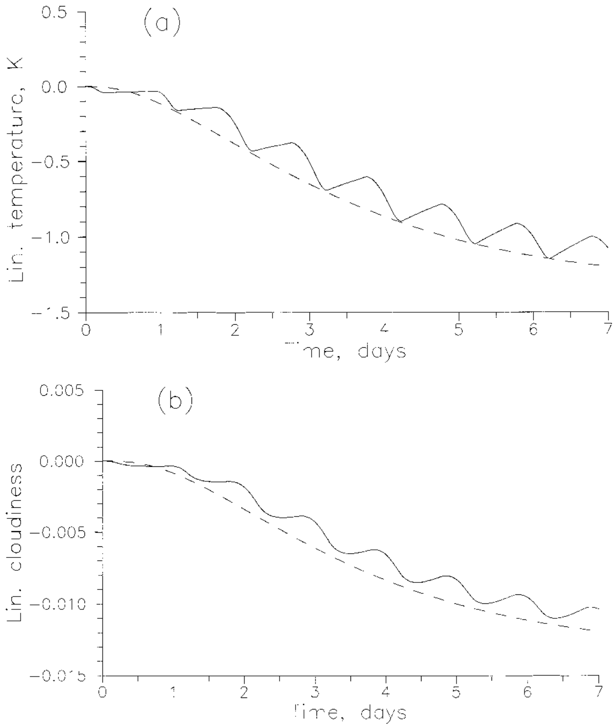

FIG. 2. Same as Fig. 1 but for the solution of the linearized forward model. Systematic deviations of the solutions for constant insolation from those for the variable insolation are due to finite value of the albedo increment $A' = 0.1$ used in the simulation.

the terrestrial atmosphere–surface system to ensure the significant diurnal variations of temperatures. The remaining parameters roughly correspond to the terrestrial atmosphere–surface system. The observables simulated were the averages of temperature and cloudiness over the last day of the total period of integration:

$$\mathbf{R} = \begin{pmatrix} \langle T \rangle \\ \langle n \rangle \end{pmatrix}. \tag{69}$$

The single model parameter varied in these numerical experiments was the albedo A. For simplicity, its basic value and its variations were kept constant over time. Thus, sensitivities of observables to this parameter reduce to partial derivatives $\partial\langle T\rangle/\partial A$ and $\partial\langle n\rangle/\partial A$. The matrix of these partial derivatives is obtained from the corresponding matrix of variational derivatives:

$$\frac{\partial \mathbf{R}}{\partial \mathbf{a}} = \int_{t_0}^{t_1} \frac{\delta \mathbf{R}}{\delta \mathbf{a}(t)}\, dt. \tag{70}$$

This expression can be derived by substitution of an arbitrary constant variation $\delta\mathbf{a}(t) \equiv d\mathbf{a}$ into the expression for the variation $\delta\mathbf{R}$, Eq. (40) which becomes a differential $d\mathbf{R}$:

$$d\mathbf{R} = \int_{t_0}^{t_1} \frac{\delta \mathbf{R}}{\delta \mathbf{a}(t)}\, d\mathbf{a}\, dt = \left[\int_{t_0}^{t_1} \frac{\delta \mathbf{R}}{\delta \mathbf{a}(t)}\, dt\right] d\mathbf{a}$$

$$= \frac{\partial \mathbf{R}}{\partial \mathbf{a}}\, d\mathbf{a}. \tag{71}$$

With arbitrary $d\mathbf{a}$, the last equality of Eq. (71) results into Eq. (70).

Figure 1 shows the results obtained for the nonlinear model with variable and constant insolation. The observables simulated for the case of variable insolation have the values

$$\begin{pmatrix} \langle T \rangle \\ \langle n \rangle \end{pmatrix} = \begin{pmatrix} 294.0 \\ 0.207 \end{pmatrix}.$$

Figure 2 shows the solution of the linearized forward problem. The increment $A' = 0.1$ was used for the linearization. Figure 3 shows the matrix solution of the corresponding adjoint problem. The values of linearized observables computed for the case of variable insolation using the solutions of both linearized forward and adjoint problems are given below:

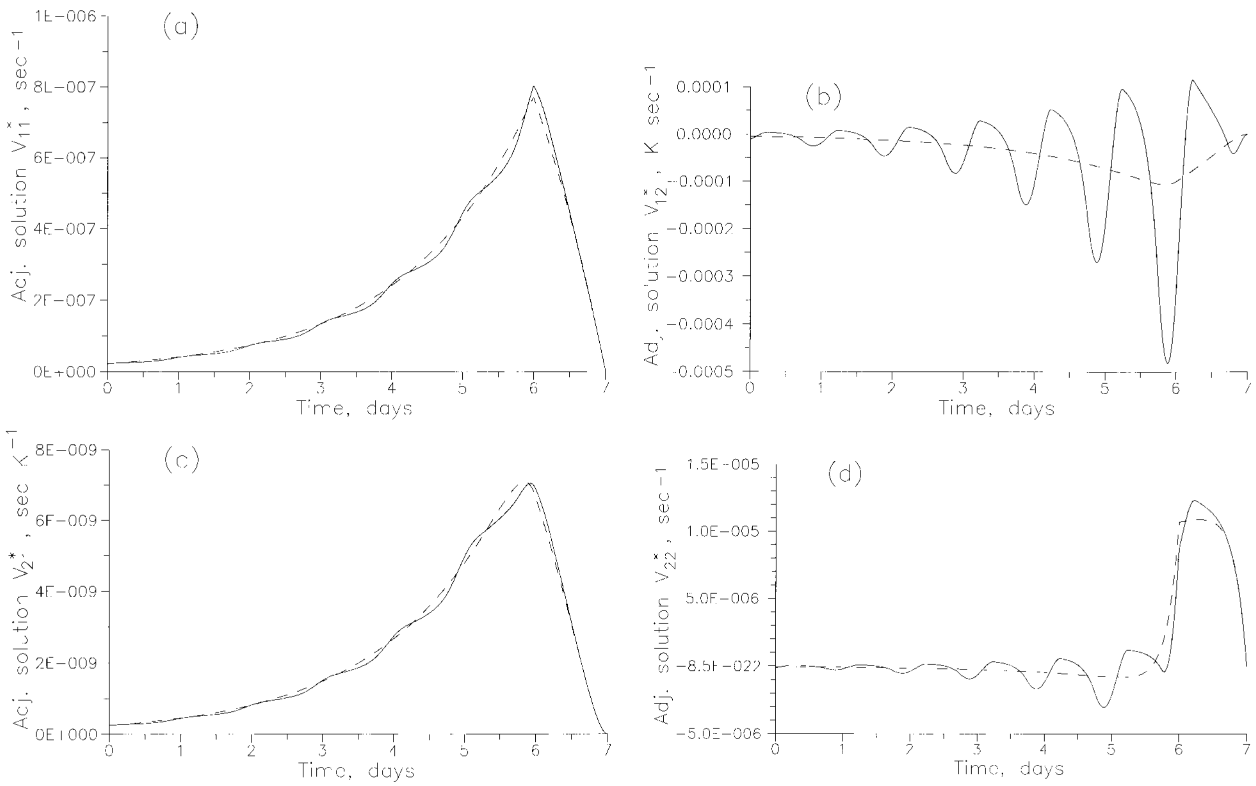

FIG. 3. Matrix solution of the adjoint model for the variable (solid line) and constant (dashed line) insolation; (a)–(d) the separate matrix elements of the solution $X^*_{jk}(t)$ (j, k = 1, 2).

$$\begin{pmatrix} \langle T' \rangle \\ \langle n' \rangle \end{pmatrix} = \begin{pmatrix} -1.06 \\ -0.0105 \end{pmatrix}.$$

The differences between the results obtained from the solutions of the linearized forward system and of the adjoint system due to numerical integration errors are within the accuracy of values presented.

Sensitivities for the basic nonlinear solutions were evaluated using the solution of the adjoint problem. The values for the case of variable insolation are presented below:

$$\begin{pmatrix} \partial \langle T \rangle / \partial A \\ \partial \langle n \rangle / \partial A \end{pmatrix} = \begin{pmatrix} -10.6 \\ -0.105 \end{pmatrix}.$$

These values can be compared with the linearized observables obtained above for $A' = 0.1$. As it can be expected, they differ by factor of $1/A' = 10$.

7. Discussion and conclusions

As demonstrated in the previous sections, the adjoint sensitivity analysis can be naturally extended to the case of multiple observables. In this case, the adjoint solution becomes a matrix with dimensions corresponding to the numbers of field variables and observables. Reducing it to a vector (one-column matrix) forces us to deal with a single observable, usually a composite one, constructed from multiple observables, like the "distance" function between the modeled and observed data. It can be anticipated that the procedure of direct construction of the adjoint operator corresponding to that of the linearized forward problem as presented in section 3 will also be applicable to the models with one or more spatial arguments where the boundary conditions will become necessary. The need for such boundary conditions is especially clear in regional models where they should be specified along the boundary of respective area. The 1D models are a logical first step here. Unfortunately, the most evident candidate, the radiative–convective model is not amenable to the straightforward linearization due to distinctly different behavior before and after onset of convection. More simple 1D candidates could be the zonally averaged energy balance models.

The variational data assimilation in meteorology using adjoint models (Talagrand and Courtier 1987) can be another area of application of this matrix approach. Here, a close analogy can be pointed out between variational data assimilation and atmospheric remote sensing. In atmospheric remote sensing, the computation of weighing functions, which represent the sensitivities of individual observables to profiles of atmospheric parameters, are routinely used to select the most informative spectral intervals and, correspondingly, the most informative observables. Similarly, in variational data assimilation, the evaluation of sensitivities of individual observables to the parameters of the atmospheric model can help in selection of the most sensitive, most informative observables.

In conclusion, the comparison of the matrix approach and the traditional vector approach to the adjoint sensitivity analysis of multiple observables can be briefly summarized as follows. The matrix of sensitivities of all individual observables with respect to the model parameters is obtained at once if the matrix approach is used. This matrix of sensitivities has to be constructed, row by row, if the traditional vector approach is used. Combined together, the vectors of sensitivities of individual observables yield the same matrix of sensitivities. Both vector and matrix approaches, when applied to the same problem, yield identical results. However, the matrix approach can provide substantial savings of computer time.

Acknowledgments. The author wishes to express his gratitude to Prof. Peter Gierasch (Cornell University) for valuable help and fruitful discussions at the early stages of this work. Also, the author is thankful to the reviewers for constructive suggestions and comments regarding the manuscript. This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

REFERENCES

Cacuci, D. G., 1981a: Sensitivity theory for nonlinear systems. I. Nonlinear functional analysis approach. *J. Math. Phys.,* **22,** 2794–2802.

——, 1981b: Sensitivity theory for nonlinear systems. II. Extensions to additional classes of responses. *J. Math. Phys.,* **22,** 2803–2812.

——, and M. D. G. Hall, 1984: Efficient estimation of feedback effects with application to climate models. *J. Atmos. Sci.,* **41,** 2063–2068.

Hall, M. C. G., 1986: Application of adjoint sensitivity theory to an atmospheric general circulation model. *J. Atmos. Sci.,* **43,** 2644–2651.

——, and D. G. Cacuci, 1983: Physical interpretation of the adjoint functions for sensitivity analysis of atmospheric models. *J. Atmos. Sci.,* **40,** 2537–2545.

——, ——, and M. E. Schlesinger, 1982: Sensitivity analysis of a radiative-convective model by the adjoint method. *J. Atmos. Sci.,* **39,** 2038–2050.

Kaminski, T., M. Heimann, and R. Giering, 1999a: A coarse grid three-dimensional global inverse model of the atmospheric transport. 1. Adjoint model and Jacobian matrix. *J. Geophys. Res.,* **104,** 18 535–18 553.

——, ——, and ——, 1999b: A coarse grid three-dimensional global inverse model of the atmospheric transport. 2. Inversion of the transport of CO_2 in the 1980's. *J. Geophys. Res.,* **104,** 18 555–18 581.

Li, Z. J., I. M. Navon, and Y. Q. Zhu, 2000: Performance of 4D-Var with different strategies for the use of adjoint physics with the FCU global spectral model. *Mon. Wea. Rev.,* **128,** 668–688.

Marchuk, G. I., 1964: Equation for the value of information from weather satellites and formulation of inverse problems. *Cosmic Res.,* **2,** 394–408.

——, 1974: Osnovnye i sopryazhennye uravneniya dinamiki atmosfery i okeana. *Meteor. i Gidrol.,* **2,** 9–37.

——, 1975a: Formulation of the theory of perturbations for complicated models. Part I: The estimation of the climate change. *Geofis. Int.,* **15,** 103–156.

——, 1975b: Formulation of the theory of perturbations for complicated models. Part II: Weather prediction. *Geofis. Int.,* **15,** 103–156.

Talagrand, O., and P. Courtier, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. I. Theory. *Quart. J. Roy. Meteor. Soc.,* **113,** 1311–1328.

Trenberth, K. E., Ed., 1992: *Climate System Modeling.* Cambridge University Press, 788 pp.

Ustinov, E. A., 1991: Inverse problem of the photometry of the solar radiation reflected by an optically thick planetary atmosphere: Mathematical framework and weighting functions of the linearised inverse problem. *Cosmic Res.,* **29,** 519–532.

——, 1992: Inverse problem of the photometry of solar radiation reflected by an optically thick planetary atmosphere. 3. Remote sounding of minor gaseous constituents and atmospheric aerosol. *Cosmic Res.,* **30,** 170–181.

——, 2000: A variational approach for computing weighting functions for remote sensing of the atmosphere in the thermal microwave spectral region. *J. Quant. Spectrosc. Radiat. Transfer,* **64,** 457–465.

——, 2001: Adjoint sensitivity analysis of radiative transfer equation: Temperature and gas mixing ratio weighting functions for remote sensing of scattering atmospheres in thermal IR. *J. Quant. Spectrosc. Radiat. Transfer,* **68,** 195–211.

Vukicevic, T., and P. Hess, 2000: Analysis of tropospheric transport in the Pacific basin using the adjoint technique. *J. Geophys. Res.,* **105,** 7213–7230.